

Tendência de Cibersegurança: IA Agêntica Requer Supervisão do Programa

14 de janeiro de 2026 - ID G00840675 - 13 min de leitura

Por: Jeremy D'Hoinne, Craig Porter

Iniciativas: Estratégia, Riscos e Oportunidades

A proliferação não gerenciada de agentes de IA está expandindo a superfície de ataque por meio da shadow AI e de automações descontroladas (*rogue automations*). Os líderes de cibersegurança devem implementar uma supervisão de segurança em camadas e adaptar seus programas para priorizar a redução de riscos com base na autonomia dos agentes e na sensibilidade dos dados.

Insights em Resumo

A proliferação rápida e não autorizada de automações de IA expande a superfície de ataque corporativa, aumentando o risco de automações de IA descontroladas. Esses **agentes de shadow AI** assumem múltiplas formas: incorporados a softwares corporativos existentes ou criados por funcionários que aproveitam os avanços tecnológicos recentes e o “**vibe coding**” para aumentar a produtividade.

Os CISOs não podem interromper essas iniciativas, já que CEOs e CIOs continuam pressionando por uma adoção agressiva da IA para manter a competitividade. Eles devem implementar rapidamente uma supervisão do programa de segurança em camadas, priorizando a redução de riscos com base no risco de negócios, determinado pela sensibilidade dos dados que a IA pode acessar e pelo escopo de suas ações autônomas.

Principais ações para CISOs:

- **Ampliar as medidas de descoberta existentes** para inventariar agentes de shadow AI fornecidos como SaaS e aqueles incorporados a softwares corporativos, agentes independentes criados com plataformas *no-code/low-code* ou desenvolvidos sob medida pela organização para apoiar os negócios.
- **Ajustar o investimento em cibersegurança:** concentrar os requisitos do início do ciclo em segurança de dados e **gestão de acessos**, reconhecendo que a maioria dos pilotos personalizados de **IA generativa** nunca chega à produção.
- **Categorizar os agentes de IA** usando dimensões simples, como a sensibilidade dos dados acessados e o nível de autonomia do agente.

- Revisar as práticas de avaliação da postura de cibersegurança para auditar elementos-chave dos agentes de IA, incluindo autenticação e detecção de anomalias comportamentais, a fim de limitar a agência excessiva.
- Criar novos playbooks de resposta a incidentes para detectar e interromper rapidamente automações de IA descontroladas, adaptando os fluxos de trabalho existentes a automações de alta frequência.

Premissas de Planejamento Estratégico

Até 2029, mais de 50% dos ataques cibernéticos bem-sucedidos contra agentes de IA explorarão problemas de controle de acesso, usando injeção de prompt direta ou indireta como vetor de ataque.

Até 2027, os custos para as empresas decorrentes de abusos de agentes de IA orientados a tarefas serão pelo menos 4 vezes maiores do que os decorrentes de **sistemas multiagentes**.

Até 2028, os CISOs e CIOs que colaborarem com os líderes de negócios para implementar um programa estruturado de cibersegurança para **IA agêntica** acelerarão em 20% as iniciativas de IA de alta agência e reduzirão os incidentes críticos em mais de 50%.

Contexto da Questão

Esta pesquisa concentrou-se na proteção do uso da IA agêntica, uma das nove tendências destacadas em Principais Tendências em Cibersegurança – 2S26.

O uso de agentes de IA está se disseminando entre funcionários e desenvolvedores, criando novas superfícies de ataque por meio da shadow AI. As plataformas no-code/low-code e o vibe coding expandem ainda mais essa superfície de ataque, impulsionando a proliferação não gerenciada de agentes de IA, código não seguro e possíveis violações de requisitos de conformidade regulatória. Muitas organizações estão buscando produtividade e redução de custos em antecipação à incerteza econômica. Como resultado, os funcionários não esperam permissão para:

- Usar automações de IA incorporadas a softwares corporativos existentes para gerar relatórios ou executar tarefas repetitivas.
- Criar automações personalizadas por meio de vibe coding ou de uma plataforma no-code/low-code
- Integrar recursos de terceiros e locais, conectando-os aos softwares corporativos que eles utilizam (por meio de protocolos como o MCP), expondo ativos da empresa.

De fato, 32% dos profissionais de TI que usam ferramentas de IA generativa no trabalho afirmam mantê-las ocultas, o que dificulta a descoberta por parte das equipes de cibersegurança. ¹

Simultaneamente, muitas iniciativas de negócios que utilizam IA generativa não estão atendendo às expectativas. Os líderes de negócios e de aplicações esperam que a automação finalmente entregue o valor prometido, já que 79% dos líderes de TI esperam que os agentes de IA proporcionem benefícios significativos de produtividade.² Eles convertem às pressas chatbots de IA personalizados em agentes de IA, conectando-os a APIs e ferramentas internas, o que acrescenta novos riscos, como agência excessiva ou automações descontroladas, à longa lista de superfícies de ataque da IA generativa. Isso significa que os CISOs não podem desacelerar nem pausar essas iniciativas, mesmo que apenas 14% dos líderes de TI pesquisados estejam muito confiantes de que seus dados/conteúdos estão prontos para a IA, tanto para interações humanas quanto de IA.³

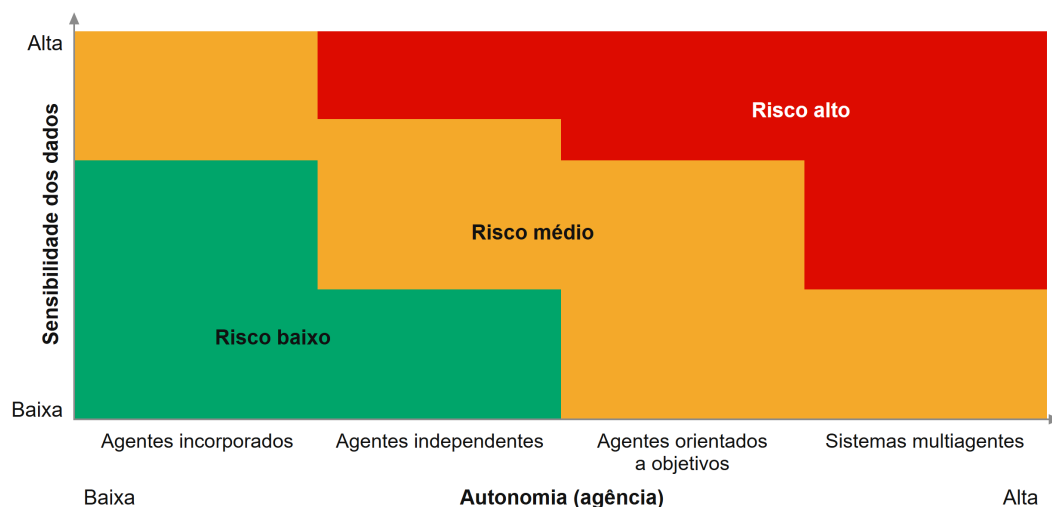
Os líderes de cibersegurança que implementarem uma estrutura adequadamente organizada para mapear os agentes de IA com base em dimensões de risco de negócios obterão a supervisão necessária para proteger e apoiar as iniciativas estratégicas de IA de sua organização.

A diversidade de agentes de IA impede que as equipes de cibersegurança criem um esquema de classificação abrangente. Uma abordagem prática é definir um número limitado de categorias com base em duas características críticas (veja a Figura 1):

- **A sensibilidade dos dados acessados pelo agente:** essa não é uma representação perfeita de todas as dimensões do risco relacionado aos negócios (“O processo é crítico?” seria outra dimensão possível), mas começar pelos dados é uma abordagem familiar e mais viável.
- **O nível de autonomia ou “agência” dos agentes de IA,** que combina o modelo de implantação (por exemplo, incorporado a uma aplicação corporativa), e que limita o escopo do agente, mas também seus principais componentes, como o tipo de modelo, que viabiliza fluxos de trabalho “orientados a objetivos”.

Figura 1: Os CISOs devem priorizar a redução de riscos para agentes de IA

Priorizar a redução de riscos para agentes de IA



Fonte: Gartner
840675

Gartner®

As equipes podem então mapear o alcance das ações que um agente de IA pode executar:

- Agentes incorporados, como parte de uma aplicação corporativa (por exemplo, AgentForce no Salesforce). A aplicação na qual estão inseridos define os dados que o agente pode acessar e oferece um enquadramento para suas ações. Ela também define as possibilidades de remediação (controle de acesso, hardening da configuração).
- Agentes independentes podem ser consumidos como SaaS ou implantados localmente. A sensibilidade dos dados depende do recurso acessado, enquanto a agência se baseia nas ferramentas disponíveis. As possibilidades de remediação estão vinculadas ao envolvimento da organização no desenvolvimento (por exemplo, plataforma no-code, low-code)
- Agentes orientados a objetivos designam agentes de IA que utilizam modelos não determinísticos, como LLMs, para gerar lógica de negócios (por exemplo, Manus ou ChatGPT Agent). O benefício de uma categoria dedicada a esses agentes é enfatizar a necessidade de práticas e controles de cibersegurança específicos voltados aos modelos de IA.
- Os sistemas multiagentes (MAS) apresentam diversas variantes. Em formato de pacote (por exemplo, uma aplicação composta por vários módulos de agentes), podem ser equiparados a um agente independente. No entanto, quando vários agentes se comunicam pela rede, os riscos e as responsabilidades variam muito em comparação com um agente independente, exigindo uma abordagem diferenciada.

A rápida adoção e a baixa maturidade dos agentes de IA e das tecnologias de cibersegurança estão forçando os CISOs a repensar seu programa de cibersegurança, redefinir as prioridades das iniciativas de segurança de IA e avaliar os benefícios e riscos para a prática de cibersegurança.

Resumo do Impacto

Apesar da baixa maturidade das tecnologias de suporte, a pressão para adquirir e adotar agentes de IA e a facilidade de acesso a ferramentas de *vibe coding* para desenvolver automações de IA são grandes demais para que a maioria das organizações resista. Os CISOs precisam recuperar o atraso rapidamente, à medida que aumentam os indícios de adoção não gerenciada de agentes de IA:

- Na pesquisa *Cybersecurity Innovation in AI Risk Management and Use* do Gartner, 61% dos profissionais seniores de cibersegurança observaram a implantação de automação por agentes de IA a partir de softwares corporativos já aprovados.
- 59% deles suspeitam ou têm evidências do uso de automação não autorizada por agentes de IA por parte dos funcionários.
- Em uma pesquisa separada, 70% dos líderes de TI que utilizam o Microsoft Copilot estão preocupados com a proliferação descontrolada de agentes do Copilot. ⁴

No entanto, as equipes de cibersegurança devem dosar cuidadosamente seus esforços e seu envolvimento de acordo com o tipo de agente de IA. Em média, apenas 41% dos protótipos independentes de GenAI chegam à produção. ⁵ Isso significa que esforços iniciais, que consomem muitos recursos, para proteger todos os pilotos de agentes de IA desenvolvidos sob medida podem acabar se revelando um investimento desperdiçado.

Implicações

Para gerenciar os riscos das implantações não gerenciadas de agentes de IA, os CISOs precisam:

Descobrir e depois inventariar agentes de shadow AI e agentes de IA autorizados

As automações de IA agêntica são fáceis de criar com ferramentas de linguagem natural e fáceis de serem mantidas por um ator mal-intencionado de forma persistente e sem detecção, devido aos desafios de visibilidade e monitoramento mencionados anteriormente. O risco de automação descontrolada também é maior porque altera a escala da perda de dados, uma vez que a extração é automatizada. Os riscos de uso indevido também são elevados, pois a tecnologia e sua implementação são novas.

Os líderes de cibersegurança devem adaptar seus programas de cibersegurança, incluindo a gestão da postura, a resposta a incidentes e a resiliência, para atender aos novos requisitos de um mundo semiautomatizado:

- Comece pelas aplicações corporativas de alto risco que possam passar a incorporar agentes de IA. Uma descoberta abrangente pode até ser possível, mas rapidamente se tornará esmagadora.
- Implemente ferramentas de descoberta que vão além da simples visibilidade das ferramentas de IA e que também possam descobrir a intenção de uso dos agentes de IA, como [Innovation Insight for AI Usage Control](#), **gateways de IA** ou plataformas mais amplas de segurança de IA. Esse processo de descoberta deve abranger tanto os sistemas de IA aprovados quanto os não aprovados, consumidos pelos funcionários, e os agentes de IA desenvolvidos internamente sob medida.
- Aproveite as funcionalidades de avaliação de riscos da ferramenta de descoberta de IA para documentar a função, o escopo, a autoridade e as permissões de cada agente.
- Crie perfis de risco dos agentes de IA com base em seu nível de autonomia, na potencial exposição de dados e na criticidade de seus processos.

Historicamente, as equipes de cibersegurança têm enfrentado dificuldades com o registro (logging) e o monitoramento de automações devido ao custo e à velocidade. Os agentes de IA acrescentam a complexidade da “automação orientada a objetivos”. Invista em ferramentas avançadas de registro para processar saídas (outputs) em linguagem natural e monitorar o comportamento dinâmico dos agentes de IA. Elabore um processo decisório para determinar se e quando esses logs devem ser centralizados em uma ferramenta de **gerenciamento de informações e eventos de segurança (SIEM)** para fornecer alertas em tempo real sobre comportamentos anômalos dos agentes.

Adaptar os programas de cibersegurança a um mundo cada vez mais automatizado

As equipes de cibersegurança enfrentam atualmente um duplo desafio com os agentes de IA: de um lado, os benefícios de utilizar a IA para aprimorar a resposta a incidentes e as operações de segurança em geral; de outro, os riscos significativos introduzidos pela adoção rápida e não gerenciada de agentes de IA.

Revise as estruturas de gestão da postura de cibersegurança para incorporar a avaliação do risco dos agentes de IA, incluindo critérios como:

- Opções de hardening para automações de IA incorporadas
- Orientações atualizadas, alinhadas à política de cibersegurança, para usuários que criam agentes de IA
- Inventário de identidades, autenticação e autorização de agentes de IA
- Práticas de desenvolvimento seguro que levem em conta o código gerado por IA
- Controle do acesso automatizado e aplicação de políticas para limitar a agência

- Guardrails e testes adversariais para agentes de IA desenvolvidos sob medida
- Monitoramento contínuo do comportamento dos agentes

Implemente medidas para aumentar a resiliência e a continuidade diante dos riscos da automação por IA. Isso inclui defesas em camadas para possíveis lacunas ou falhas na **governança de IA**, como:

- Monitoramento ampliado para detectar automações novas e descontroladas
- Validação e testes dos comportamentos dos agentes de IA
- Integração de supervisão humana (human-in-the-loop) para funções críticas

Atualize os processos de resposta a incidentes de cibersegurança para definir como as equipes podem responder à atividade de agentes de IA descontrolados. Os planos e processos devem abordar:

- Como garantir a visibilidade de todas as automações no escopo do programa
- Como detectar comportamentos anormais
- Como identificar e interromper agentes de IA comprometidos
- Como adaptar os fluxos de trabalho de resposta a incidentes à velocidade dos processos automatizados

Priorizar as iniciativas de segurança de IA com base no risco de negócios e na maturidade tecnológica

Planeje e aloque recursos para primeiramente proteger os agentes de IA de maior risco, definidos como aqueles incorporados a sistemas sensíveis, que acessam dados sensíveis e que são implementados ou implantados com maior autonomia. Por meio de avaliações internas de risco, as organizações podem determinar quais agentes têm maior probabilidade de causar interrupções caso sejam comprometidos.

Os líderes de cibersegurança devem concentrar os requisitos do início do ciclo em ações que seriam difíceis ou caras de implementar em estágios posteriores, começando pela segurança de dados e pela gestão de acessos. Em seguida, devem colaborar com as equipes que autorizam ou implementam automações de IA para determinar o alcance das ações de cibersegurança e o seu ritmo. Em muitas organizações, apenas uma pequena porcentagem dos pilotos chegará à produção. Isso pode levar a um enorme desperdício de energia e a conflitos entre equipes. Por fim, eles precisam implementar controles em tempo de execução e testar sua eficácia como parte de um processo padrão quando uma aplicação ou um agente entra em produção.

Sobre as Principais Tendências de Cibersegurança do Gartner para 2026

Esta tendência é uma das nossas Principais Tendências de Cibersegurança para 2026. São as tendências que consideramos mais relevantes e impactantes. Nossas tendências se enquadram em três temas principais:

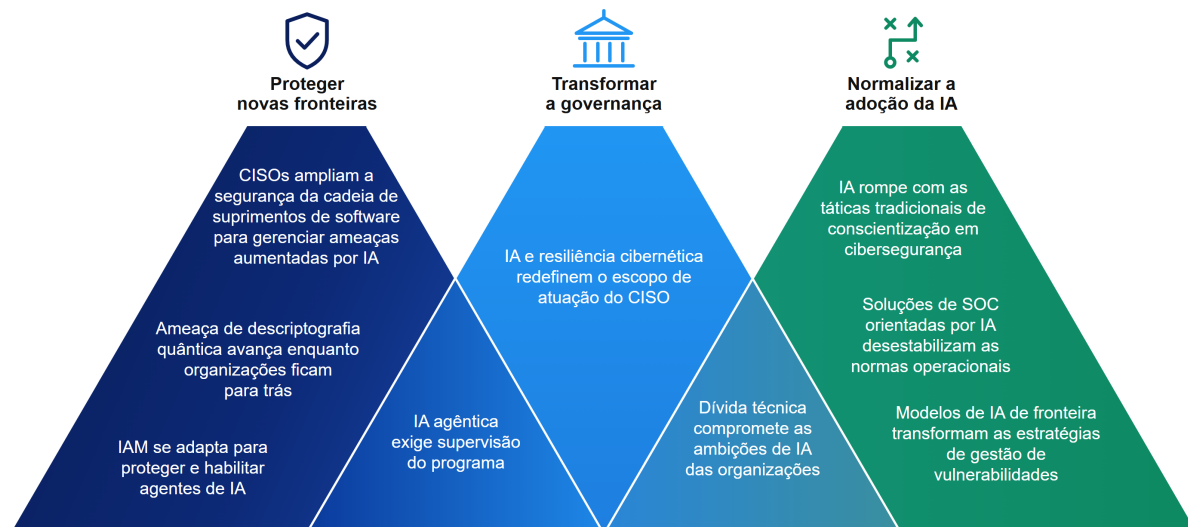
- Transformar a governança
- Proteger novas fronteiras
- Normalizar a adoção da IA

Essas tendências de cibersegurança não existem isoladamente — elas se interconectam (veja a Figura 2), e várias se enquadram em mais de um tema. A importância potencial das tendências para sua organização varia de acordo com a maturidade organizacional, mas também com o setor, as necessidades de negócios e os planos estratégicos elaborados anteriormente.

Trabalhe com outros executivos para avaliar os impactos e benefícios de nossas tendências. Isso permitirá determinar quais tendências individuais — ou qual combinação estratégica — terão o impacto mais significativo em sua organização e no ecossistema em que ela opera. Examine o potencial das tendências em relação à situação específica de sua organização, incorpore-as ao seu planejamento estratégico para os próximos anos e ajuste seus modelos de negócios e operações de forma adequada.

Figura 2: Principais Tendências de Cibersegurança para 2026

Principais tendências de cibersegurança para 2026



Fonte: Gartner
860263

Gartner®

Colaboradores

Alex Michaels, Chiara Girardi, Arthur Sivanathan, Will Candrick

Evidências

¹ 2025 Technology at Work Report, Ivanti.

² **Pesquisa *Generative and Agentic AI in Enterprise Applications* do Gartner de 2025.** Este estudo foi realizado para compreender os principais desafios e oportunidades na implantação de ferramentas de IA generativa (GenAI) e onde as organizações devem concentrar seus investimentos em IA. A pesquisa também teve como objetivo entender em que estágio as organizações se encontram em sua jornada com agentes de IA e o que pensam sobre esses agentes. A pesquisa foi realizada on-line de maio a junho de 2025 com 360 entrevistados de organizações com pelo menos 250 funcionários em tempo integral, de todos os setores (exceto software de TI), na América do Norte (n = 149), Europa (n = 140) e Ásia/Pacífico (n = 71). Foram estabelecidas cotas flexíveis por país, porte da empresa, tipo de função e nível hierárquico do entrevistado para garantir uma boa representatividade da amostra. As organizações deveriam ter implantado, ou planejar implantar em menos de um ano, pelo menos uma ferramenta de IA generativa em pelo menos um domínio central de aplicações corporativas: aplicações de ambiente de trabalho digital, aplicações de gestão do relacionamento com o cliente ou aplicações de planejamento de recursos empresariais. Os entrevistados ocupavam cargos de liderança de equipe ou superiores, excluindo o nível C, e estavam envolvidos na implementação de ferramentas de IA generativa; foi exigido que tivessem determinadas responsabilidades em relação a essas ferramentas. Aviso: os resultados desta pesquisa não representam conclusões globais nem o mercado como um todo, mas refletem a percepção dos entrevistados e das empresas pesquisadas.

³ **Pesquisa *Cybersecurity Innovations in AI Risk Management and Use* do Gartner de 2025.** Esta pesquisa foi realizada para compreender como as organizações estão gerenciando os riscos de cibersegurança da IA generativa (GenAI) e das técnicas de IA que a sustentam. A pesquisa foi realizada on-line de 21 de março a 9 de maio de 2025 com 302 líderes de cibersegurança nas regiões da América do Norte (n = 181), EMEA (n = 71) e Ásia/Pacífico (n = 50). As organizações qualificadas declararam receita corporativa de pelo menos US\$ 250 milhões, ou equivalente, no ano fiscal de 2024, e os entrevistados pertenciam à alta gestão de cibersegurança e estavam envolvidos em atividades relacionadas à gestão de riscos de cibersegurança de IA em sua organização. Aviso: os resultados desta pesquisa não representam conclusões globais nem o mercado como um todo, mas refletem a percepção dos entrevistados e das empresas pesquisadas.

⁴ **Pesquisa Microsoft 365 and Copilot do Gartner de 2025.** Esta pesquisa foi realizada on-line de 10 de março a 1º de abril de 2025 para compreender como as organizações estão governando, utilizando e dando suporte ao Microsoft 365. Também teve como objetivo determinar a adoção e a utilização do Microsoft 365 Copilot e seu impacto na posição do Microsoft 365 na organização. Participaram 215 líderes de TI que influenciam ou tomam decisões relacionadas ao Microsoft 365. Cento e setenta e sete concluíram a pesquisa — 112 do *Research Circle* do Gartner (um painel gerenciado pelo Gartner) e 65 contatados por meio da divulgação do link da pesquisa em conferências do Gartner e de ações junto a clientes. Os entrevistados que informaram sua localização estavam na América do Norte (n = 96), EMEA (n = 75), Ásia/Pacífico (n = 9) e América Latina (n = 8); 27 entrevistados não indicaram sua localização. Aviso: os resultados desta enquete não representam conclusões globais nem o mercado como um todo, mas refletem a percepção dos entrevistados e das empresas consultadas.

⁵ **Pesquisa AI Mandates for the Enterprise do Gartner de 2024.** Este estudo foi realizado para compreender como a IA e a IA generativa (GenAI) estão sendo adotadas pelas empresas, com foco em áreas como estratégia de IA, dados, governança, letramento, engenharia, organização, portfólio e valor, para ajudar os clientes a acompanhar a rápida evolução da IA. A pesquisa foi realizada on-line de outubro a dezembro de 2024 com 432 entrevistados dos EUA (n = 181), Reino Unido (n = 70), França (n = 50), Alemanha (n = 50), Índia (n = 51) e Japão (n = 30). Foram estabelecidas cotas por porte de empresa e por setor para garantir uma boa representatividade da amostra. As organizações deveriam ter implantado pelo menos um caso de uso de IA em produção. Os entrevistados foram selecionados entre executivos de nível C (por exemplo, chief AI officer, chief data officer, **chief data scientist**, chief digital officer, chief information officer, chief operating officer, chief technology officer ou equivalente) ou cargos acima de vice-presidente. Todos os entrevistados deveriam ter alto envolvimento em pelo menos uma iniciativa de IA. Aviso: os resultados desta pesquisa não representam conclusões globais nem o mercado como um todo, mas refletem a percepção dos entrevistados e das empresas pesquisadas.

Recomendado pelos autores

Alguns documentos podem não estar disponíveis em sua assinatura atual do Gartner.

[How to Secure Custom-Built AI Agents](#)

[How to Secure Enterprise Agentic AI Ambition](#)

© 2026 Gartner, Inc. e/ou suas afiliadas. Todos os direitos reservados. Gartner é marca registrada da Gartner, Inc. e de suas afiliadas. Esta publicação não pode ser reproduzida ou distribuída de nenhuma forma sem a permissão prévia e por escrito do Gartner. Ela consiste nas opiniões da *Business and Technology Insights Organization* do Gartner, que não devem ser interpretadas como declarações de fato. Embora as informações contidas nesta publicação tenham sido obtidas de fontes consideradas confiáveis, o Gartner se isenta de todas as garantias quanto à exatidão, integridade ou adequação dessas informações. Embora os insights do Gartner possam abordar questões jurídicas e financeiras, o Gartner não fornece aconselhamento jurídico ou de investimentos, e seus insights não devem ser interpretados ou utilizados como tal. Seu acesso e uso desta publicação são regidos pela [Política de Uso do Gartner](#). O Gartner se orgulha de sua reputação de independência e objetividade. Seus insights são produzidos de forma independente pela *Business and Technology Insights Organization*, sem contribuição ou influência de terceiros. Para mais informações, consulte "[Princípios Orientadores de Independência e Objetividade](#)". Os insights do Gartner não podem ser utilizados como dados para o treinamento ou o desenvolvimento de inteligência artificial generativa, aprendizado de máquina, algoritmos, software ou tecnologias relacionadas.